MD

18

UNITED NATIONS

ECONOMIC AND SOCIAL COUNCIL

Distr.
GENERAL

E/ECA/PSD.5/47
28 December 1987

Original: ENGLISH

ECONOMIC COMMISSION FOR AFRICA

Fifth session of the Joint Conference
of African Planners, Statisticians
and Demographers

Addis Ababa, Ethiopia, 21 - 28 March 1987

LC/[illegible]381

REPORT OF THE TRAINING WORKSHOP ON THE USE OF SOFTWARE PACKAGES IN STATISTICAL COMPUTING FOR ENGLISH-SPEAKING GOVERNMENT STATISTICIANS AND PLANNERS ADDIS ABABA, 21 SEPTEMBER - 2 OCTOBER 1987

TABLE OF CONTENTS

	Paragraphs	Pages
I. ATTENDANCE AND ORGANIZATION OF WORK		
(i) Opening and duration of meeting	1 - 2	1
(ii) Attendance	3	1
(iii) Agenda	4	1
II. ACCOUNT OF PROCEEDINGS		
Use of Statistical Software Packages	5 - 9	2
Introduction to Micro-Computers and Disk Operating System (DOS commands)	10 - 14	3
Presentation of selected software packages on micro-computers:		6
PC-COXTALLY - Summary cross-tabulation	15 - 21	6
EASWESPOP - Data entry	22 - 25	7
EASWESPOP - Fertility estimate programs	26 - 28	8
PICK - Multi-user operating system and data base	29 - 36	8
IMPS - Integrated Micro-computer Processing System	37 - 41	10
LOTUS 1-2-3	42 - 44	11
DBASE III plus	45 - 47	11
Essential features of selected software packages: A comparative presentation of Biomedical programmes (BMDP), the Statistical Analysis System (SAS), and the Statistical Packages for the Social Sciences (SPSS)	48 - 51	12
Introduction to and examples on the use of basic instructions of SPSS/PC+ and practical exercises	52 - 83	13
ANNEX I		

I. ATTENDANCE AND ORGANIZATION OF WORK

(i) Opening and duration of meeting

1. The Training Workshop on the Use of Software Packages in Statistical Computing for English-speaking Government Statisticians and Planners was held at ECA headquarters, Addis Ababa from 21 September to 2 October 1987. It was organized by the Statistics Division of the United Nations Economic Commission for Africa (UNECA) in collaboration with the Pan African Documentation and Information System,(PADIS), and the Population Division. The Commonwealth Fund for Technical Cooperation (CFTC) financed the cost of the participation of a number of officials from Commonwealth African countries. The US Bureau of the Census and the United Nations Headquarters through the Department of Technical Cooperation for Development (DTCD) provided teaching materials and lecturers.

2. The Training Workshop was opened by Prof. Adebayo Adedeji, Executive Secretary of the United Nations Economic Commission for Africa.

(ii) Attendance

3. The Training Workshop was attended by 10 specialists responsible for statistical data processing from the following countries: Ethiopia, Gambia, Kenya, Lesotho, Malawi, Mauritius, Seychelles, Sierra Leone, and Swaziland.

(iii) Agenda

4. The following is the programme for the workshop:

1. Opening address
2. Use of statistical software packages in statistical computing
3. Introduction to micro-computers and Disc Operating System (DOS)
4. Presentation of selected software packages on micro-computers:

4.1	PC-COXTALLY	- Summary cross-tabulation
4.2	EASWESPOP	- Data entery - Fertility estimate programs
4.3	PICK	- Multi-user operating system and data base
4.4	IMPS	- Integrated Micro-computer Processing system
4.5	LOTUS 1-2-3	- Spread sheet
4.6	DBASE III	- Data Base Management System

5. Essential features of selected software packages: A comparative presentation of Biomedical Computer Programmes (BMDP), the Statistical Analysis System (SAS), and the Statistical Package for the Social Sciences (SPSS)

6. Introduction to and examples on the use of SPSS/PC and practical exercises.

II. ACCOUNT OF PROCEEDINGS

Use of Statistical Software Packages in Statistical Computing

(Agenda item 2)

5. A representative of the secretariat presented the paper "Use of Statistical Software Packages in Statistical Computing", ECA/STAT/SPSC/SM/87/2. He pointed out that the production of statistical information in the form of tables, reports, and graphs is an important contribution to socio-economic development planning, policy formulation and decision making. The rapid technological advancement in computer hardware/software has brought the use of computers to the disposal of professionals involved in statistical computing without the need for the assistance of computer experts.

6. Statistical software packages can cover all the steps required in the production and analysis of statistical information, from the collection of the raw data to the final presentation of the results. The main uses of statistical software packages for data collection, editing, file handling, tabulation, analysis and graphical presentation were discussed at length. It was stressed that understanding the underlying statistical techniques and their proper interpretation was of great importance in statistical data analysis.

7. Some of the most widely used statistical software packages with mainframes, mini and micro computers, their features, uses and abuses were also presented. The use of micro computers the hardware and software of which could be acquired at low cost were also discussed. Micro computers have proved to be efficient tools for the collection, analysis and reporting of statistical data. They are also used for the collection of large volumes of census and survey data with many video display units connected to them to serve as data capture work-stations. Micro computers are easy to use using English like language by non programers. They assist professionals in their daily work by reducing the laborious job of manual computation, analysis and presentation of statistical data.

8. The problem of training in the use of statistical software packages was raised by the participants. Software packages for micro computers have, step by step, instruction in their installation and use. Most of them have on-line help facilities, some even have tutorials. It is, however, necessary to have some training specially for mainframes and mini computers.

9. The lack of training, sufficiently detailed installation and running instructions for CENTS 4 and CONCOR packages developed by the US Bureau of the Census was raised. The acquisition of micro computers and software packages were the other problems raised by the participants. A representative of the US Bureau of the Census confirmed the possibility of organizing regional workshops regularly. He stated that he will present all the questions raised at the seminar to the office of the US Bureau of the Census. A representative from the United Nations Department of Technical Cooperation for Development (DTCD) stated that the United Nations could help in acquiring software packages developed by the United Nations free of charge. Commercial packages could also be obtained through the United Nations at the lowest possible price. A bilateral cooperation program might be used to get assistance for the acquiring mini or micro-computer.

Introduction to the Use of Micro-Computers and Disc Operating System (DOS Commands)

(Agenda item 3)

I. Micro-computers

10. During the presentation of this subject the following items were discussed:

i). Types of micro-computers. It was specified that the following standards of machines exist:

(a) Micro-computers with the 8088/86 Central Processing Unit (CPU); typical machines: IBM PC, IBM PC/XT;

(b) Micro-computers with the 80286 CPU; typical machine: IBM PC/AT;

(c) Micro-computers with the 80386 CPU (the latest development of CPU's by Intel; typical machines: Compaq, Deskpro 386, IBM PS/2 model 80. It was mentioned that in the micro-computer world the so called "industry standard" refers to the micro-computers of IBM, and all so-called industry standard compatibles are designed to work as the IBM machines of the different types;

ii). Micro-computer peripherals: As the most important kind of peripherals the following items were discussed: Storage units and printers. In the field of data storage the following hardware was mentioned:

(a) External diskette drives;

(b) Disc drives might be of different physical sizes: 8" diskettes are the standard for mini-computers and mainframes. This type of diskettes might be used for data transfer between mini-computers and main frames on the one hand and micro-computers on the other, without the need for sophisticated connections like modems or networks for access to mini-computer or mainframe data. 5 1/4" diskettes are the standard for the IBM PC micro-computers and compatibles. Today, two different standards for this size of diskette exist: the double sided, double density 5 1/4" diskette with a capacity of 360 Kilo Byte (to be found in the IBM PC, IBM PC/XT micro-computers and compatibles) and the high density 5 1/4" diskette with 1.2 Mega Byte capacity (to be found in the IBM PC/AT micro-computers and compatibles); 3 1/2" diskettes (the standard for so called portable micro-computers and the new IBM PS/2 systems). Again two types of capacities are today's standard: firstly, portable micro-computers and the low-end IBM PS/2 system 30 use 3 1/2" diskette of 720 Kilo Byte. Secondly, 3 1/2" diskettes with 1.44 Mega Byte capacity used with the IBM PS/2 systems 50, 60 and 80. If a statistical office has often to transfer data between different formats, it might be wise to purchase a machine which is able to transform between these formats (e.g. DiskMaker II);

(c) External hard drives: There are external hard disks available with capacities up to 512 Mega Bytes;

(d) Bernoulli boxes and tape streamers: This type of backup systems allow theoretically for unlimited storage capacity as they are removable. Even though they could be used in the same manner as a hard disk, i.e. directly accessing files on the tape or the bernoulli box, they are mostly used for backup of hard disks only. Usually they allow the storage of some 20 Mega Bytes on a single tape or Bernoulli box plater but also 40 and 60 Mega Byte versions are available.

iii). The current printer market is sub-divided into the following type of printers:

(a) Daisy wheel printers: Daisy wheel printers usually provide the user with a 96 character set. This means that the user has no access to the full IBM character set of 255 characters. Some producers offer special wheels containing semi-graphics, but this means that the wheel has to be changed while printing. Graphics output, as produced by LOTUS 1-2-3, cannot be printed using a daisy wheel printer. On the other hand, daisy wheel printers can be used to produce stencils as the pressure of the letters is the same as with normal type writers.

(b) Matrix-/Inkjet printers: The main advantage of matrix-/inkjet printers is the facility to print graphics as well as text. In addition this type of printers are relatively cheap. Today's matrix printers with up to 2 pins on the printing head allow for near letter quality printing and will do for most of the usual typing in statistical offices;

(c) Laser printers: The most recent development of relatively cheap laser printers allow high quality printing for micro-computer users. This type of printers might be used for publications and reports, to produce camera ready originals.

II. Disc Operating System (DOS)

11. The most widely used operating system on micro-computers is the M(icro) S(oft) - D(isk) O(perating) S(ystem). The PC-DOS, used with the IBM PC's is created by the same company and has the same functions as MS-DOS, so for the user there is no difference between these systems.

12. The MS-DOS is composed of internal and external commands. Internal commands (such as DIR, CD RD, etc.) are at the disposal of the user at any time after loading the operating system. Other commands (like FORMAT) are files contained on the operating system diskettes which must be executed like any other executable files

13. MS-DOS accepts only specific files as executable files from the command line: command files (with the extension .COM), executable files (with the extension .EXE) and batch files (with the extension .BAT).

14. Command files and executable files contain machine code which can be executed immediately. Batch files can contain, apart from some special batch file commands, any sort of allowed MS-DOS commands; i.e. with a batch file the user can execute internal and external MS-DOS commands as well as any other command or executable or batch file. If wanted the user can free himself from typing many different commands from the MS-DOS command line by creating a batch file and only executing once this batch file.

Presentation of selected software packages on micro-computers

(Agenda item 4)

PC-COXTALLY - Summary cross-tabulation

(Sub-item 4.1)

15. A historic background of the development of the package noted that the package was first developed for mini computers and written in RPG-II and was called XTALLY. Originally the package was intended for small computers such as IBM System 32, System 34, System 36 and others in the IBM series as well as other brands of computers. Due to the more widespread popularity of COBOL in many countries as compared to that of RPG, a UN field expert converted XTALLY to COBOL and the new name of COXTALLY was assigned. XTALLY and COXTALLY have been successfully used for processing (tabulation) of national censuses in various countries. The PC version has now also been in use for several years and has been successfully used for processing of census data in small countries.

16. COXTALLY and its PC version may have some limitations as regards headings, titles and stubs as well as flexibilities as compared to other tabulation packages such as CENTS-4. However, there are some specific advantages which cannot be overlooked:

- It is intended for non-programmers and those who have no computer background. It is very easy to learn. An average person without any computer background can master it in a few hours;

- Many functions are automatic. Once the variables and categories have been defined, the user does not have to worry about the table format, whether it will fit in one page or not and the headings for columns and rows;

- It requires a minimum of work for producing the same table as compared to other packages. It produces fast results.

- Since the final tables produced are in ASCII (for PC) it is always possible to read or import the table into any word processor and make any changes required to make the tables of good quality for publishing;

17. In the workshop, various stages of defining variables, category sets, table definitions, execution, printing, aggregating to higher levels and percentages were presented and demonstrated. The DEMO was described.

18. The participants requested to have hands on experience with the package and thus extra half day was allowed for the practice of the PC-COXTALLY package.

19. Although it was the intention of the workshop to create an awareness among the participants about the package, thanks to the simplicity of the package, by the end of the practice session, the participants were confident of being able to use the package by themselves.

20. The presentation stimulated demand for the package. The participants were advised that they could obtain the package and its manual from the Software and Support for Population Data Processing Project (INT/83/P56).

21. The participants of the workshop were also informed of the availability of other packages, at no cost to the recipients, from the above mentioned project. For the availability of software packages and the address of the Software and Support for Population Data Processing Project, please refer to Annex I.

EASWESPOP - Data entry

(Sub-item 1.2)

22. The package provides on easy method of entering data and verifying it. It consists of three modules:

1. Layout design of the record;
2. Entry of data; and
3. Verification of data.

23. Features such as definition of field types as numeric, printable character, constant, incremental and left or right justification are included. It also provides for the filling of blanks in the field by any character. One range can be checked per field.

24. A number of most recommended features in any data entry package are lacking in this one. No inter-record consistency checks and no intra-record consistency checks are provided. The provision of multi range checks for the same field should be a desirable plus. Most of all, the verification programme would need to be strengthened further. The limitation of number of records per file for verification is not conducive to census and survey types of data. The extra work of tailoring down the files for verification would greatly delay any work.

25. The workshop participants found the package easy to work with. They agreed that notwithstanding its weaknesses and the high prices one would pay for much superior commercial packages, the package could be useful installations where no data entry package is available.

EASEWESPOP - Fertility estimate programs

26. Although the package was not included in the original agenda of the workshop, the organizers of the workshop were able to include the demonstration of the package in the already tight schedule of the workshop.

27. The participants went through the procedures of entering sample data themselves and obtained the results. They found the package easy to use and the documentation clear.

28. The availability of other demographic packages such as MORTPAK-LITE and MCPDA from the United Nations were noted.

PICK - Multi-user operating system and data base

(Sub-item 4.3)

29. The capabilities of PICK operating system and data base management system can be fully appreciated when demonstrated on the computer. In the workshop it was stated that it has been the desire to present a demonstration on the computer however, due to the unavailability of required equipment (IBM PC-AT computer) it was not possible to present any demonstration.

30. At the system operating level, it was noted that PICK was a virtual memory machine, swapping in and out of memory 512 bytes long pages called frames and provided multi user capability. Due to its unique software architecture, and special assembler level language, the virtual memory application has been implemented with the minimum of overhead on the mini and micro computers.

Thanks to the implementation of overlapping buffers, PICK has made it possible to provide multi user capability even on micro computers without any significant or noticeable reduction in response time. On IBM PC-AT, it is possible to connect 10 dummy terminals under PICK. On Fujitsu micro computers, it is possible to add up to 25 ports.

31. PICK has its own hierarchical file structure with four distinct levels. The first level being the System file, (being only one in the entire system) pointing to one or more or many dictionary files in the account or directory which inturn points to the lowest level of one or more data files. The dictionaries provide definitions of either other dictionaries of data files and items. Each item is defined by an item in the related dictionary. The structure of a record (known as item in PICK) is different from its counterpart in the conventional files systems:

- The record length (item length) is variable and can be up to 32 KB;

- Fields of records (referred to as attributes in PICK) are also of variable length and can be from 0 to 32 KB long;

- Fields can be divided into variable length values. There is no counterpart for this in conventional file structure;

- Values can also be subdivided into "sub-values" of variable length which have no equivalent in conventional systems.

32. Due to its file structure and data organization, PICK has been able to implement a true relational data base management system.

33. In fact, the PICK system was designed for data base management purposes. Together with its file structure and data organization, PICK has developed a powerful non-procedural query language called ACCESS which is English-like and easy to use.

34. The PICK software, in addition to its query language, also includes PICK/BASIC, PROC (a procedure language), and TCL (Terminal Control Language).

35. A number of other commercial software are also available for PICK which increase the power of PICK to handle spread-sheet and graphic tasks. The spread-sheet package called COMPUSHEET, in addition to LOTUS 1-2-3 like features, provides facility to read directly from other files and other spread sheet and return required data to the working spread sheet. The ACCUPLOT provides facility for preparing coloured graphs in the form of bar charts,

line graphs and pie charts. SCREENGEN and WIZARD are two data entry packages available for the software. Text processor which supports lazer-jet printers is also available.

36. It was noted that PICK was a commercial package and protected by copy-right. Software and Support for Population Data Processing project could assist and advise in the procurement of and training in the use of the software.

IMPS - Integrated Micro-computer Processing System

(Sub-item 4.4)

37. Mr. Thomas Melaney of the International Statistical Programs Centre (ISPC) of the U.S. Bureau of the Census presented a demonstration of the Integrated Microcomputer Processing System (IMPS) produced by ISPC for census and survey data processing. The demonstration began with an examination of the rationale for the development of IMPS and an explanation of the system design and objectives. The requirements of the target microcomputer configuration were then given.

38. An overview was presented of the six modular components of IMPS: data entry, editing, tabulation, statistical analysis, demographic analysis and operational control. The structure and operations of the selected software components currently being used in IMPS were investigated as well as those components which are being projected for use in the near future.

39. The following IMPS software packages were examined:

Entry point	-	One of the currently recommended choices of commercially available data entry packages (RODE-PC is the other recommended data entry system);
Data dictionary	-	A system developed by ISPC to define the structures of data file, for use in editing and tabulation (not yet available; to be released in the near future);
CONCOR 2.4	-	Data editing system developed by ISPC;
CENTS 4.2	-	Data tabulation system developed by ISPC;
PC-CARP	-	Statistical analysis package developed by the University of Iowa;

MCPDA - Microcomputer programs for demographic analysis developed by Westinghouse.

All of these packages with the exception of PC-CARP were demonstrated on the microcomputer.

40. An analysis of several case studies of countries currently using microcomputer systems for processing data from national censuses was presented. This analysis examined the problems and solutions involved in the decision to use microcomputers, the configurations chosen, software selections made, training of personnel, development of data processing programs and data processing operations.

41. Participants discussed the processing needs of national statistical offices and how the components of IMPS could be used. They were requested to contact ISPC should they want further information on obtaining any of the software used in IMPS.

LOTUS 1-2-3

(Sub-item 4.5)

42. This spread sheet program was introduced by means of a 40 minute video film, which showed how 1-2-3 might be used by a business man supervising sales staff.

43. After this, the secretariat made a presentation of three statistical applications. The first of these highlighted the advantages of 1-2-3 for producing information in tabular form, for publication in statistical bulletins for example. He demonstrated the ease with which similar tables, of derived percentages for example, can be constructed. The second application conceived the compilation of national accounts statistics, while the third was designed to calculate a quarterly index of manufacturing production. One feature of this application was the technique employed to prevent totals being shown when the data was incomplete.

44. A second video film, on some of the more advanced features of 1-2-3, was shown after the end of the morning session.

DBASE III plus

(Sub-item 4.6)

45. A short introduction on the feature of the Data Base Management System DBASE III plus was given. It was emphasized that DBASE III plus only was selected as an example of a Data Base Management System on micro- computers and that there are other data base systems with similar features for micro-computers and that their quality may differ from that of DBASE III plus.

46. It was pointed out that DBASE III plus can be used by both programmers and non-programers. In a practical exercise the participants learned how to use the DBASE III plus utility ASSIST to create a small sample data base and a data entry screen for this data base.

47. After the practical exercise a short overview of the DBASE III plus programming languages and its practical application was given. The introduction was supplemented by a short paper, "An example of the application of DBASE III plus: The program compindx" (ECA/STAT/SM/87/6).

Essential features of selected software packages: A comparative presentation of Biomedical Computer Programmes (BMDP), the Statistical Analysis System (SAS), and the Statistical Package for the Social Sciences (SPSS)

(Agenda item 5)

48. The paper ECA/STAT/SPSC/SM/87/3 was presented by a representative of the secretariat. The statistical software packages SPSS, BMDP and SAS were presented using an example of simple data description. The commands necessary for simple data description in the three packages were shown and explained. Differences and similarities between the three packages were pointed out.

49. It was emphasized that SAS differs considerably from both BMDP and SPSS in that SAS is a programming language in its own right. This can prove useful when complicated data transformations are to be performed.

50. It was also mentioned that although SPSS may not be the most comfortable statistical software package available for micro-computers (there are also versions of SAS and BMDP for micro-computers) its availability on most mini and mainframe computers and its popularity will most likely make it the most popular statistical software package on micro-computers.

51. A participant mentioned that he would like to see a demonstration of SPSS. The representative of the secretariat stated that during the following sessions of the seminar the participants would work with SPSS in practical exercises.

Introduction to and examples on the use of basic instructions of SPSS/PC+ and practical exercises

(Agenda item 6)

52. The presentation of this agenda item was subdivided into 11 parts:

(a) General introduction to SPSS/PC+;
(b) Preparing and defining data, FREQUENCIES, CROSSTABS, DESCRIPTIVES procedures;
(c) Transformation and selection of data;
(d) Creation, restructuring and use of system file;
(e) Command description and order;
(f) Introduction to REVIEW;
(g) Submission of commands from a file;
(h) SET, SHOW and REPORT commands;
(i) SPSS/PC+ other commands;
(j) Overview of SPSS/PC+ TABLES, and
(k) Overview of SPSS/PC+ advanced statistics.

These 11 parts were presented by 4 different representatives of the secretariat.

53. When presenting the general introduction to SPSS/PC+, a representative of the secretariat recalled that SPSS/PC+ was used for the training workshop not on the grounds of preference among existing software packages but only for its availability at ECA on micro-computers. He emphasized that the objective of the training was not to make the participants experts in SPSS/PC+ but only to interest them in its use and enable them to continue the training at home by themselves. He further presented SPSS/PC+ as a command-based system and general purpose statistical software package designed to run interactively on the IBM/PC with hard disk, and at least 384 KB of RAM. The SPSS/PC+ system was composed of three main types of commands:

(a) Operation commands which provide information about or affect the way the system operates;
(b) Data definition and manipulation commands which instruct the system on how and where to read data, how to transform data, how to define missing data, how to select and weight cases, and how to provide labels and formats; and
(c) Procedure commands which cause the system to display data, provide statistics, compare groups, and provide utilities.

54. Data analysis concepts were introduced regarding particularly cases, variables and the system command syntax. Then it was shown how to enter and connect commands:

(a) Interactively;
(b) Through the command INCLUDE, and
(c) through the in-built editor REVIEW.

Entering data was also briefly discussed through use of commands DATA LIST, BEGIN DATA and END DATA, SAVE, GET, EXPORT, IMPORT and WRITE. The commands JOIN and AGGREGATE were also mentioned for restructuring files.

55. The different types of SPSS/PC+ files were mentioned as well as the order of commands illustrated with an example of a SPSS/PC+ run, the enhancement of the SPSS/PC+ basic system with advanced statistics and TABLES, and the tutorial which present easy-to-learn 14 lessons on the basic system. The participants were informed that every part of that presentation would be resumed in details in subsequent sessions on SPSS/PC+, and the installation of SPSS/PC+ on micro-computer would be practically shown to them.

56. In the discussions which followed, participants wanted to know whether ECA could copy the system for them if diskettes were sent. It was replied that this was not possible because SPSS/PC+ was a commercial software package with a copyright. A participant asked whether the tutorial was easy-to-use. It was confirmed that the one diskette tutorial provided an easy and rapid way to learn the basic system of SPSS/PC+.

57. Concerning creation, restructuring and use of system files, the following commands: SAVE, GET, JOIN, were discussed in details. Practical exercises were done by the participants thereon. Due to lack of time, the command AGGREGATE was left aside.

58. A representative of the secretariat gave further explanation on the DOS commands required to install the SPSS/PC+ software package and the participants could successfully practice its installation on one of the available micro-computers before general considerations about the system were presented.

59. Three ways for submitting commands were mentioned:

(a) Interactively;
(b) Using review editor during an SPSS/PC+ session, and
(c) Using an external command file.

60. The method of correcting errors while entering the commands by means of the backspace, backslash keys and the editing keys F1, F2, and F3 was explained.

61. The syntax diagram was then presented as well as the layout of the commands, the rule for abbreviating them and the rule for naming variables.

62. The following data definitions and manipulation commands were shown: DATA LIST, BEGIN DATA, END DATA, FORMATS, MISSING VALUE.

63. Because all the participants were interested in the well known fixed format data file, only the key word FIXED in the data list command was presented. Further explanations were given on the treatment of numeric blanks and the "Get Blanks = Value" command was shown.

64. Other commands were presented on how to exit an SPSS/PC+ session. HELP, DISPLAY, and LIST.

65. The first statistical procedure to be presented was DESCRIPTIVES.

66. All these commands have permitted the participants to do practical exercises immediately as the training was proceeding over a test data file which consisted of 2709 records of 28 demographic variables.

67. Due to the background of the participants there had been no need to give explanations on the underlying descriptive statistical techniques used by the system. The intervention of the lecturer during the practical exercise was therefore limited to reminding the participants of the syntax of commands.

68. The session on TABLES began with a general presentation of the TABLES options of SPSS/PC+. The system produces camera-ready tables in several ways:

(a) When the default SCREEN = ON is in effect;
(b) When the default LISTING = ON is in effect;
(c) When PRINTER = ON is specified on the SET command, and
(d) The content of the tables are automatically written to a special file named SPSS.TAB. PRINT TABLES commands can be used for printing the tables.

69. The attention of the participants was drawn to the management of the tables file, especially when TABLES is to be run repeatedly. In this case it is recommended to rename the file.

70. Explanations were given on how to build a table. Tables are specified in one, two or three dimensions. The stub is the first dimension and defines the rows. The banner is the second dimension and defines the columns. Layers from the third and highest dimension define a stack of individual tables. The local subcommands were presented first . The techniques for concatenating and nesting tables, for exploring percentages and other calculation functions were explained in detail.

71. Concatenation allows to present more than one table in a single display while nesting makes it possible to create a table with multiple dimensions and reduce the number of dimensions. The STATISTICS sub-command was described. It is used to specify the calculations needed such as percentages, means, variance, etc. The way of obtaining summary statistics using FTOTAL and PTOTAL sub-commands was also explained.

72. Finally the global subcommands were over viewed; when running a table. They control printing characteristics for a variety of features: page size, line drawing, spacing, highlighting, lines of text for titles and footnotes, printing of dates and page numbers.

73. The participants were very much interested in the tables options and ran many practical exercises. Unfortunately, they could not have their results printed since no printer was available.

74. The use of the INCLUDE command was explained to the participants for submitting commands from a command file. This command can be nested and it was indicated that the user was able to develop a library of commonly used sets of command which can be included in SPSS/PC+ as the first command in every session.

75. The other procedures presented were FREQUENCIES, CROSSTABS, MEANS, and PLOT. Detailed explanations were given for running each of them and practical exercises were demonstrated.

76. SPSS/PC has inbuilt procedures and commands for the transformation and selection of data. Data transformation is changing an observed data value to another, changing a range of values to one value, or generating new variables. Such transformations are sometimes required for ease of analysis, reporting and to group a set of values. The three SPSS commands used to transform data, RECODE, COMPUTE and COUNT were presented with practical exercises.

77. The RECODE command changes an alphabetic data to another alphabetic or numeric data, range of data values to one value, or to assign a specific value for the missing data using the keyword MISSING. The COMPUTE command assigns new values for a given variable or generates a new variable using arithmetic expressions. The arithmetic operators used for the expression are:

+ for addition,
- for subtraction,
* for multiplication,
/ for division, and
** for exponentiation.

SPSS has also inbuilt mathematical functions that are used along with the arithmetic operators. These functions are:

ABS	Absolute value
RND	Argument rounded to the nearest integer
TRUNC	Argument truncated to its integer part
MOD10	Remainder resulting when the argument is
SQRT	Square root
EXP	Exponential
LG10	Base 10 logarithm
LN	Natural or Naperian logarithm
ARTAN	Arctangent
SIN	Sine
COS	Cosine

Also available are the random number functions UNIFORM and NORMAL, the cross-case function LAG and the date function YRMODA.

78. The COUNT command creates a new variable which contains the number of occurences of a given data value for a set of variables in each case. The data values to be checked could be numeric or a string of alphanumeric data. It can also be used to count user-missing values.

79. SPSS has the facility of conditional data transformation using the IF command along with logical expression. The logical expression is built using the following relational operators:

EQ	Equals to
NE	Not equal to
LT	Less than
LE	Less than or equal to
GT	Greater than
GE	Greater than or equal to

80. Two or more relations could be joined using the logical operators AND and OR. The NOT logical operator is also used to reverse the true or false status of an expression.

81. The new values and variables generated by the data transformation commands can be labelled to give more description of the values and variables. This is done using the VARIABLE LABELS and VALUE LABELS and MISSING VALUE commands.

82. SPSS allows the user to select data conditionally and unconditionally, to give weight to certain variables and to sort the data. The commands used to select data are SELECT IF, PROCESS IF, SAMPLE, and the N command.

83. The SELECT IF and PROCESS IF commands select data based on arithmetic expression permanently and temporarily respectively. The SAMPLE command selects data at random using a pseudo-random-number generator. The N command is used to select the first N cases from the data. The data to be analyzed could be re-ordered using the SORT command. Data values could also be adjusted by giving different weights using the WEIGHT command.

ANNEX I

Software and Support for Population Data Processing Project (INT/83/P56)

The following statistical and demographic software is available for developing countries and international organizations.

XTALLY:	- A cross tabulation package for mainframe computers with RPG-II compiler;
COXTALLY:	- A COBOL version of XTALLY tabulation package for mini and main frame computers;
PC-COXTALLY:	- A tabulation package for micro-computers which will run on IBM PC, IBM PC-XT, IBM PC-AT and compatible micro's;
MORTPACK:	- A demographic analysis package based on indirect methods of estimating demographic indicators. The package is available both for mainframes and micro-computers. The micro computer version runs on IBM PC-XT, IBM PC-AT and compatibles;
CPDA and MCPDA:	- Demographic analysis package available both for mainframes and micro's;
EASWESPOP DATA ENTER:	- A generalized data entry package for IBM PC-XT, IBM PC-AT and comptibles;
EASWESPOP FERTILITY:	- A generalized package for estimation of fertility based on indirect techniques for IBM PC-XT, IBM PC-AT and compatibles;
FIV-FIV SIN-SIN:	- A population projection package for IBM PC-XT, IBM PC-AT and compatibles;
RAPID PROJECTION:	- A projection package for IBM PC-XT, IBM PC-AT and compatibles.

Countries or organizations interested in obtaining the above packages may write directly or through their UN funded projects or through the UNDP and/or the UNFPA resident office in their countries to the following address:

Mr. MICHAEL R. LACKNER
Coordinator, Software and Support
for Population Data Processing
Two UN Plaza, Room DC2-1570
United Nations
New York, N.Y. 10017
U.S.A